Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 56 (2016) 456 – 460

www.elsevier.com/locate/procedia

9th International Conference on Digital Enterprise Technology - DET 2016 – "Intelligent Manufacturing in the Knowledge Economy Era

Forecasting of Intraday Interval Arrivals for Small and Medium sized Call Centers with Emergencies

Lan Xu*, Xiuwu Hu, Xiu-li Wang, Guodong Huang

School of Economics and Management, Nanjing University of Science and Technology, Nanjing 210094, China

* Corresponding author. Tel.: 15850567955; *E-mail address:* xulan1309@126.com

Abstract

In order to improve service level and responsiveness, many market-oriented enterprises gradually pay attention to the role of call centers. Call center supported by digital technology can provide data to help these enterprises in terms of knowledge mining, data analysis, and even intelligent manufacturing. This paper focuses on the forecast method of call center, presenting a two-sector model. The upper model proposed is used to identify the characteristics of drastic incoming calls change based on recent and historical data, and the bottom model proposed consists of several prediction models that process different data characteristics.

Peer-review under responsibility of the scientific committee of the 5th CIRP Global Web Conference Research and Innovation for Future Production

Keywords: Call Center, Interval Call Center Arrivals, Two-Sector Prediction Model, Singular Value Decomposition, Emergency

1. Introduction

Call center supported by digital technology can provide decision-making information for production enterprises, and it is also an important channel between production enterprises and market. Whether enterprises built call center by themselves or cooperate with the third party call center, the service level of call center can be equivalent to the service quality of the production enterprise for customers in market. Hence, enhancing the operating capability of call center can improve service quality of production enterprises and response to customer's need. For most call centers, 60-70% of operating expenses are capacity costs, e especially in human resource costs. Optimizing staff scheduling and adjusting their work schedules dynamically can improve customer service levels and save labor costs. The service skills of staffs (or telephone operators) in call center can also be enhanced through reasonable arrangement of staffs work the prerequisite of optimizing the allocation of staff resources is that the forecast accuracy of interval call arrivals is increased for different incoming calls. Therefore establishing a forecast models of interval call arrivals and efficient algorithms are needed for call centers.

Call arrivals denote the number of telephone calls when the staff receives from customers in a fixed period. In recent years, domestic and foreign researchers are interested in forecasting issues of call arrivals. Shen and Huang (2008) [1] used singular value decomposition technique to construct a forecasting model of interval call arrivals, particularly through the North American financial companies call centers experimental analysis of sample data to verify the proposed model predicts better than traditional and some previous findings. Li and Xin (2009) [2] used the principle of least squares support vector machine to establish time series forecasting models of daily call arrivals and interval call arrivals. Weinberg et al. (2007) [3], Aktekin and Soyer (2011) [4] described the discrete-time Bayesian state-space model of non-homogeneous Poisson process used to forecast call arrivals. Du and Jiang (2012) [5] analyzed the historical data of a large call center, and used auto-regressive integrated moving average model prediction

This work was supported by the National Natural Science Foundation of China under [grant number 71571101] and [grant number 71171114].

Peer-review under responsibility of the scientific committee of the 5th CIRP Global Web Conference Research and Innovation for Future Production
doi:10.1016/j.procir.2016.10.089

model to forecast the interval call arrivals. Ibrahim and Ecuyer (2013) [6] considered working day and data correlation effects between interval arrivals and proposed fixed-effects, mixed-effects, and bivariate models to predict interval arrivals. The experimental results showed that the binary mixed-effects model had a better effect than the fixed or single variable mixing.

According to our knowledge, there are many studies of forecasting call arrivals for large call centers, in which the number of incoming call is bigger than small call center, and the interval call arrivals has little fluctuation in the same business day. For small and medium call center, in most cases we find that the numbers of incoming call within interval periods are stable, but sometimes the call volume will sudden increase.. Experimental calculations show that using the existing prediction model is not enough. In order to get more accurately predict on interval call arrivals for small and medium call centers with emergencies, a two-sector prediction model of interval call arrivals is proposed.

2. Two-Sector Prediction Model

Intraday interval arrivals of small and medium sized call centers with emergencies have three features. (1) In steady state, the same interval arrivals of the same weekday in different weeks have significant correlation. (2) The same interval arrivals of continuous weekdays also have a certain correlation in the same week and continuous interval arrivals on the same day. (3) When emergencies happen, the correlation of the same interval arrivals of continuous weekdays in the same week and the correlation of continuous interval arrivals on the same day strengthen. The essential idea of two-sector prediction model is to distinguish the feature information of drastic data change possibly and then use different historical data to forecast.

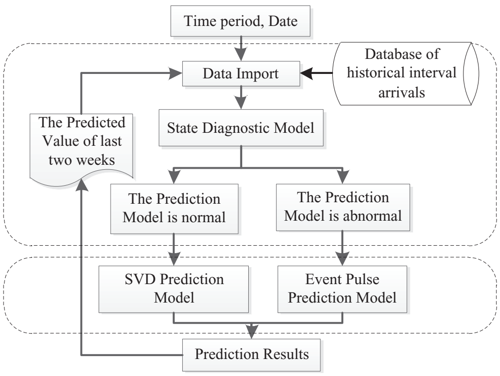

Fig. 1. the structure of two-sector prediction model.

Figure 1 shows the structure of two-sector prediction model proposed. To get information indicator that possible changes drastically, the diagnostic layer is based on analysis and comparison of actual historical interval arrivals and predictive values calculated under the assumptions of steady state. The information indicators are as the inputs of interval arrivals forecasting model. According to the information indicators, the forecasting model at the bottom gives the predictive values of interval arrivals with different mathematical models and different historical data.

Some symbols of the model have the following meanings:

I : The number of the predicted week;

K : Days of the week;

J : Time periods of interval arrivals predicted;

$x_{i,k,j}$: The actual value of interval arrivals in one time period, k=1,2,$\cdots$,7, j=1,2,$\cdots$,m;

$x_{i,k}$: The actual value of the intraday call volume profile of the k th day of the i th week, $x_{i,k} = \left(x_{i,k,1}, x_{i,k,2}, \cdots\cdots, x_{i,k,m}\right)$;

X_k : An $n \times m$ matrix that records the call volumes for the m time periods in the k th day of the week in n weeks, $x_k = \left(x_{i,k,j}\right)_{n\times m}$;

$\hat{x}_{i,k,j}$: The predictive value of interval arrivals in one time period;

$\bar{x}_{i,k} = \left(\bar{x}_{i,k,j}\right)_{1\times m}$, $\bar{x}_{i,k,j} = \frac{1}{n}\sum_{l=i-n}^{i-1} x_{l,k,j}$;

$D_{i,k} = \sum_{j=1}^{m} x_{i,k,j}$;

$\bar{D}_{i,k} = \frac{1}{n}\sum_{l=i-n+1}^{i} D_{l,k}$.

m_0 : Threshold value in advance, dividing the time into two categories of m_0 time period before tor after.

3. The Diagnostic Layer of Two-Sector Prediction Model

For forecasting of interval arrivals of small and medium sized call centers which may be affected by unexpected events, the state diagnostic model has a crucial role. In the case, the state diagnostic model filters out interference of random factors through the setting threshold. The closer actual data is, the more likely it implies the future trend information.

The algorithm of status diagnosis model:

Step 0: Set time periods m , weeks n , warning thresholds a , b and c , interval threshold m_0 , lower limit lb and upper limit ub of intermediate variables $\eta_{i,k,j}$.

Step 1: Calculate the interval arrivals by the singular value decomposition predictive model.

Step 2: Calculate intermediate variables $\alpha_{i,k} = \frac{D_{i,k}}{\bar{D}_{i,k}}$,

$$\beta_{i,k} = COUNTIF\left(x_{i,k}, "\gamma_{i,k,j} \geq a"\right),\ \gamma_{i,k,j} = \frac{x_{i,k,j}}{\bar{x}_{i,k,j}} \text{ and}$$

$$\eta_{i,k,j} = \frac{\hat{x}_{i,k,j}}{x_{i,k,j}}, (j = 1,2,\cdots,m,\quad c < m).$$

Step 3: Diagnose whether interval arrivals are abnormal on the same day of week in the last week and the week before last week, or not.

$$C_I = [c_{I1}, c_{I2}],\qquad c_{I1}, c_{I2} \in \{0,1\}, \tag{1}$$

$$c_{I1} = \begin{cases} 1 & ,\alpha_{I\text{-}1,K} \geq a, \beta_{I-1,K} \geq b \\ 0 & ,others \end{cases},$$

(2)

$$c_{I2} = \begin{cases} 1 & ,\alpha_{I\text{-}2,K} \geq a, \beta_{I-2,K} \geq b \\ 0 & ,others \end{cases}.$$

(3)

When $\alpha_{I\text{-}1,k} \geq a$ and $\beta_{I-1,k} \geq b$, we judge that interval arrivals are abnormal on the same day of week in last week, as c_{I1}=1 . When $\alpha_{I\text{-}2,k} \geq a$ and $\beta_{I-2,k} \geq b$, we judge that interval arrivals are abnormal on the same day of week in the week before last week, as c_{I2}=1 . According to the results of C_I , it gives the matrix of actual historical interval arrivals.

Step 4：Diagnose whether interval arrivals are abnormal on that day and the day before that day, or not.

When $J \leq m_0$, we determine whether forecasting model is abnormal according to the interval arrivals in the same time period before that day and the interval arrivals of one time period before the time period on that day. When $J > m_0$, we determine whether forecasting model is abnormal according to the interval arrivals in the same time period before that day and the interval arrivals of two time periods before the time period on that day. Therefore, two state diagnostic models are as follows relying on whether $J \leq m_0$ or not.

(1)When the number of the time period is $J \leq m_0$,

$$L_{I,K,J} = \begin{cases} 1 & ,\eta_{I,K\text{-}1,J} \notin [lb,ub]\, and \quad \eta_{I,K,J\text{-}1} \notin [lb,ub] \\ 0 & ,others \end{cases}.$$

(4)

If $\eta_{I,K\text{-}1,J}, \eta_{I,K,J\text{-}1} \notin [lb,ub]$, it means the abnormal event for the interval arrivals in the same time period before the previous day and the interval arrivals of one time period before the time period on the previous day . In other words, the predicted values of interval arrivals by the prediction model are irregular, if $L_{I,K,J} = 1$. On the contrary, it means that the predicted values of interval arrivals are normal, if $L_{I,K,J} = 0$.

(2)When the number of the time period is $J > m_0$,

$$L_{I,K,J} = \begin{cases} 1 & ,\eta_{I,K\text{-}1,J}, \eta_{I,K,J\text{-}1}, \eta_{I,K,J\text{-}2} \notin [lb,ub] \\ 0 & ,others \end{cases}.$$

(5)

If $\eta_{I,K\text{-}1,J}, \eta_{I,K,J\text{-}1}, \eta_{I,K,J\text{-}2}$ are all beyond the range of $[lb,ub]$, it means that the interval arrival in the same time period before the previous day and the interval arrivals of two time periods before the time period on the previous day are both abnormal. In other words, the predicted values of interval arrivals by the prediction model are irregular, if $L_{I,K,J} = 1$. On the contrary, it means that the predicted values of interval arrivals are normal, if $L_{I,K,J} = 0$.

4. The Prediction Layer of Two-Sector Prediction Model

According to characteristic index information that the state diagnostic model exports in section 3, we use an appropriate forecasting model in the prediction layer. Singular value decomposition prediction model forecast model and event pulse prediction model are specifically described below.

4.1. Singular Value Decomposition Prediction Model

The thought of the singular value decomposition prediction model: At first the dimensionality of historical matrix of interval arrivals is reduced by matrix singular value decomposition; then we use the autoregressive model to forecast interval arrivals. Singular value decomposition gives an effective rank of the matrix without changing the metric characteristics of data, and gives the best approximation matrix rank reduction by minimizing the prediction error. Dimensionality reduction, parameter estimation and reconstruction constitute singular value decomposition prediction model, and they are introduced following.

4.1.1 Dimensionality reduction by singular value decomposition

Firstly historical data of interval arrivals in n weeks is divided into seven segments of the time series matrix, according to different day of the week. Then interval arrivals can be expressed as

$$X_k = \begin{bmatrix} x_{1,k,1} & x_{1,k,2} & \cdots & x_{1,k,m} \\ x_{2,k,1} & x_{2,k,2} & \cdots & x_{2,\mathrm{k},m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,k,1} & x_{n,k,2} & \cdots & x_{n,k,m} \end{bmatrix}. \tag{6}$$

X_k is an $n \times m$ matrix that records the interval arrivals on the k th day for *n* weeks, with each day having m time periods. If k is given, the rows and columns of X_k correspond respectively to weeks and time periods within a day. X_k is abbreviated as $x_i = (x_{i1}, x_{i2}, \cdots, x_{im})^T$ after subscript omitted.

The matrices $x_1, x_2, x_3, \cdots$ constitute a vector time series. Let $f_1, f_2, \cdots, f_P$ represent vector bases, where $f_p = (f_{1p}, \cdots, f_{mp})^T$, $p = 1, 2, \cdots, P$. We consider the following decomposition,

$$x_i = \beta_{i1} f_1 + \beta_{i2} f_2 + \cdots + \beta_{iP} f_P + \varepsilon_i,$$

(7)

where ε_i is the error term. We choose the appropriate parameters $\beta_{i1}, \beta_{i2}, \cdots, \beta_{iP}$ and basis vectors $f_1, f_2, \cdots, f_P$ by making the error terms small in magnitude,

$$\min_{\beta_{i1},\cdots,\beta_{iP}; f_1 \cdots f_P} \| \varepsilon_i \|^2$$

(8)

The SVD of the matrix X can be expressed as

$$X = USV^T,$$

(9)

where an $n \times m$ matrix $U = (u_1, u_2, \cdots, u_m)$ and an $m \times m$

matrix $V=(v_1,v_2,\cdots,v_m)$ are orthogonal. Diagonal matrix $S=diag(s_1,s_2,\cdots,s_m)$, $s_1\ge s_2\ge\cdots\ge s_m$. It then follows from (8) that

$$x_i=s_1u_{i1}v_1+s_2u_{i2}v_2+\cdots+s_ru_{ir}v_r\,, \tag{10}$$

with the largest P singular values, we have the following approximation:

$$x_i\cong s_1u_{i1}v_1+s_2u_{i2}v_2+\cdots+s_Pu_{iP}v_P\,. \tag{11}$$

We take $\beta_{ip}=u_{ip}$ and $f_p=s_pv_p$.

4.1.2 Parameter Estimation and Reconstruction of Prediction Model

(1)Without using the data of interval arrivals on the previous day

Forecast interval arrivals according to historical data of interval arrivals on the same day of the week in n weeks before. Through time series $\{\beta_{i1}\},\{\beta_{i2}\},\cdots,\{\beta_{iP}\}$, with first order autoregressive model, we can forecast coefficients $\hat\beta^{TS}_{n+h,1},\hat\beta^{TS}_{n+h,2},\cdots,\hat\beta^{TS}_{n+h,P}$ with first order autoregressive model. Then a forecast of x_{n+h} is given by

$$\hat x^{TS}_{n+h}=\hat\beta^{TS}_{n+h,1}f_1+\hat\beta^{TS}_{n+h,2}f_2+\cdots+\hat\beta^{TS}_{n+h,P}f_P \tag{12}$$

(2)With using the data of interval arrivals on the previous day

More information can be used to forecast interval arrivals on that day. We forecast interval arrivals on that day relying on historical data in n weeks before and interval arrivals in m_0 time periods before on that day.

Let $x^e_{n+1}=(x_{n+1,1},\cdots,x_{n+1,m_0})^T$ represents the interval arrivals in m_0 time periods before on that day; let $x^l_{n+1}=(x_{n+1,m_0+1},\cdots,x_{n+1,m})^T$ represents the interval arrivals that will be forecasted; let

$$F^e=\begin{pmatrix} f_{11} & f_{12} & \cdots & f_{1P}\\ f_{21} & f_{22} & \cdots & f_{2P}\\ \cdots & \cdots & \cdots & \cdots\\ f_{m_0,1} & f_{m_0,2} & \cdots & f_{m_0,P}\end{pmatrix}. \tag{13}$$

Minimize the following formula to estimate the parameters $\beta_{n+1,1},\beta_{n+1,2},\beta_{n+1,P}$

$$\min_{\beta_{n+1,1},\cdots,\beta_{n+1,P}}\{\sum_{j=1}^{m_0}|x_{n+1,j}-(\beta_{n+1,1}f_{j1}+\cdots+\beta_{n+1,P}f_{jP})|^2+\lambda\sum_{p=1}^{P}|\beta_{n+1,p}-\hat\beta^{TS}_{n+1,p}|^2\}.$$

It is also expressed as minimization,

$$\min_{\beta_{n+1}}\{(x^e_{n+1}-F^e\beta_{n+1})^T(x^e_{n+1}-F^e\beta_{n+1})+\lambda(\beta_{n+1}-\hat\beta^{TS}_{n+1})^T(\beta_{n+1}-\hat\beta^{TS}_{n+1})\} \tag{14}$$

To calculate from (14) can get estimated values of the parameters $\beta_{n+1,1},\beta_{n+1,2},\cdots,\beta_{n+1,P}$,

$$\hat\beta^{PLS}_{n+1}=(F^{eT}F^e+\lambda I)^{-1}(F^{eT}x^e_{n+1}+\lambda\hat\beta^{TS}_{n+1})\,. \tag{15}$$

The estimated value of x^l_{n+1} may be given by the following formula

$$\hat x^{PLS}_{n+1,j}=\hat\beta^{PLS}_{n+1,1}f_{j1}+\hat\beta^{PLS}_{n+1,2}f_{j2}+\cdots+\hat\beta^{PLS}_{n+1,P}f_{jP} \tag{16}$$

where $j=m_0+1,\cdots,m$, the parameter λ depends on the feature of the data .

4.2. Event Pulse Prediction Model

According to the result that state diagnostic model identifies possible exception, the prediction layer utilizes the event pulse prediction model to forecast the interval arrivals. The event pulse prediction model doesn't depend on historical data accumulated for a long time, but it relies on the data in recent two weeks, especially the closest interval arrivals. The event pulse prediction model is described below.

Intraday arrivals and historical data information form variables $S_{I,K,J}$ and $\mu_{I,K,J}$ .The variable $S_{I,K,J}$ reflects the extending trend of intraday arrivals. As a rate, the variable $\mu_{I,K,J}$ reflects structure information in the history matrix. The model takes the trend information and structural information into account to forecast the interval arrivals.

We calculate the predicted value of the interval arrivals $\hat x_{I,K,J}$ by following formula:

$$\hat x_{I,K,J}=S_{I,K,J}\times\mu_{I,K,J}\,, \tag{17}$$

where

$$S_{I,K,J}=\begin{cases}\sum_{j=1}^{J-1}x_{I,K,j} & ,J\le m_0\\ \sum_{j=J-P}^{J-1}x_{I,K,j} & ,J>m_0\end{cases}, \tag{18}$$

$$\mu_{I,K,J}=\begin{cases}\dfrac{\bar x_{I,K,J}}{\sum_{j=1}^{J-1}\bar x_{I,K,j}} & ,J\le m_0\\[2ex] \dfrac{\bar x_{I,K,J}}{\sum_{j=J-P}^{J-1}\bar x_{I,K,j}} & ,J>m_0\end{cases}. \tag{19}$$

5. Conclusion

In this paper, we propose a two-sector model to forecast the number of incoming calls in a small and medium sized call center. The two-sector model with the state diagnostic has a strong adaptive ability and has higher forecast accuracy for interval arrivals of call centers with emergencies. The model proposed is feasible in real-time dynamic forecasting with digital enterprise technology, since the model proposed is a data driven and has a fast calculating speed.

References

[1] Shen H, Huang JZ. Interday forecasting and intraday updating of call center arrivals[J]. Manufacturing & Service Operations Management, 2008, 10(3): 391-410.

[2] Li DC, Xin ZH. Large call-centers' arrival-rates prediction models based on the least squares support vector machine[J]. Control Theory & Applications; 2009, (7): 815-818.

[3] Weinberg J., Brown L.D., Stroud J.R. Bayesian forecasting of an inhomogeneous Poisson process with applications to call center data[J].

Journal of the American Statistical Association; 2007, 102(480): 1185-1198.
[4] Aktekin T., Soyer R.. Call center arrival modeling: A Bayesian state-space approach[J]. Naval Research Logistics; 2011, 58(1): 28-42.
[5] Du SG, Jiang F. Forecasting model of phone calls amount for large-scale call centres and its analyses[J].Computer Applications and Softwar; 2012, 29(9):126-128.
[6] Ibrahim R, L'Ecuyer P. Forecasting call center arrivals: Fixed-effects, mixed-effects, and bivariate models[J].Manufacturing & Service Operations Management; 2013, 15(1): 72-85.